

CRF Processing Date:  
 Edited by:  
 Verified by:

Serial Number:

09/751,299A

☐

Changed a file from non-ASCII to ASCII

☐

Changed the margins in cases where the sequence text was "wrapped" down to the next line

☐

Edited a format error in the Current Application Data section, specifically:

☐

Edited the Current Application Data section with the actual current number. The number inputted by the applicant was ☐ the prior application data; or ☐ other

☐

Added the mandatory heading and subheadings for "Current Application Data".

☐

Edited the "Number of Sequences" field. The applicant spelled out a number instead of using an integer.

☐

Changed the spelling of a mandatory field (the headings or subheadings), specifically:

☐

Corrected the SEQ ID NO when obviously incorrect. The sequence numbers that were edited were:

☐

Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited:

☐

Corrected subheading placement. All responses must be on the same line as each subheading. If the applicant placed a response below the subheading, this was moved to its appropriate place.

☐

Inserted colons after headings/subheadings. Headings edited included:

☐

Deleted extra, invalid, headings used by an applicant, specifically:

☐

Deleted: ☐ non-ASCII "garbage" at the beginning/end of files; ☐ secretary initials/filename at end of file; ☐ page numbers throughout text; ☐ other invalid text, such as

☒

Inserted mandatory headings, specifically: <220>, sequences # 2, 4

☐

Corrected an obvious error in the response, specifically:

☐

Edited identifiers where upper case is used but lower case is required, or vice versa.

☐

Corrected an error in the Number of Sequences field, specifically:

☐

A "Hard Page Break" code was inserted by the applicant. All occurrences had to be deleted.

☐

Deleted **ending** stop codon in amino acid sequences and adjusted the "(A)Length:" field accordingly (error due to a PatentIn bug). Sequences corrected:

☐

Other:

\*Examiner: The above corrections must be communicated to the applicant in the first Office Action. DO NOT send a copy of this form.

3/1/95

6/29/01 OIPE 16  
 #  
 RECEIVED  
 DEC 3 2001  
 TECH CENTER 600/290

OIPE

## RAW SEQUENCE LISTING

DATE: 06/29/2001

PATENT APPLICATION: US/09/751,299A

TIME: 14:23:08

Input Set : A:\DIVER1440-2final.txt

Output Set: N:\CRF3\06292001\I751299A.raw

```

3 <110> APPLICANT: Madden, Mark
4     Weiner, David P.
5     Chaplin, Jennifer A.
7 <120> TITLE OF INVENTION: METHODS FOR PRODUCING ENANTIOMERICALLY PURE
8     ALPHA-SUBSTITUTED CARBOXYLIC ACIDS
10 <130> FILE REFERENCE: DIVER1440-2
12 <140> CURRENT APPLICATION NUMBER: US 09/751,299A
13 <141> CURRENT FILING DATE: 2000-12-28
15 <150> PRIOR APPLICATION NUMBER: 60/254,414
16 <151> PRIOR FILING DATE: 2000-12-07
18 <150> PRIOR APPLICATION NUMBER: 60/173,609
19 <151> PRIOR FILING DATE: 1999-12-29
21 <160> NUMBER OF SEQ ID NOS: 4
23 <170> SOFTWARE: PatentIn Ver. 2.1
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 1041
27 <212> TYPE: DNA
28 <213> ORGANISM: Unknown Organism
30 <220> FEATURE:
31 <223> OTHER INFORMATION: Description of Unknown Organism: Obtained from an
32     environmental sample
34 <220> FEATURE:
35 <221> NAME/KEY: CDS
36 <222> LOCATION: (1)..(1041)
38 <400> SEQUENCE: 1
39 atg tcg gag ccc atg acg aag tat cgc ggc gcg gcg gtg cag gcc gcg 48
40 Met Ser Glu Pro Met Thr Lys Tyr Arg Gly Ala Ala Val Gln Ala Ala
41 1 5 10 15
43 ccg gtg ttc ctc gat ctc gac cgc aca gtc gag aaa gcg atc ggc ctg 96
44 Pro Val Phe Leu Asp Leu Asp Arg Thr Val Glu Lys Ala Ile Gly Leu
45 20 25 30
47 atc gag cag gcg gcc aag cag gac gtg cgc ctg atc gca ttc cca gag 144
48 Ile Glu Gln Ala Ala Lys Gln Asp Val Arg Leu Ile Ala Phe Pro Glu
49 35 40 45
51 act tgg att ccc ggc tat ccc ttt tgg ata tgg ctg ggc gcg ccg gct 192
52 Thr Trp Ile Pro Gly Tyr Pro Phe Trp Ile Trp Leu Gly Ala Pro Ala
53 50 55 60
55 tgg ggc atg cgc ttc gtc cag cgc tat ttc gag aat tcg ctc gtg cgc 240
56 Trp Gly Met Arg Phe Val Gln Arg Tyr Phe Glu Asn Ser Leu Val Arg
57 65 70 75 80
59 ggc agc aag cag tgg cag gcc ctg gcg gat gcg gcc cgc cgc cac ggc 288
60 Gly Ser Lys Gln Trp Gln Ala Leu Ala Asp Ala Ala Arg Arg His Gly
61 85 90 95
63 atg cat gtc gtg gcc ggc tat agc gag cgc gcg ggc ggc agc ctc tat 336
64 Met His Val Val Ala Gly Tyr Ser Glu Arg Ala Gly Gly Ser Leu Tyr
65 100 105 110
67 atg ggc cag gcg atc ttc ggc ccc gat ggc gat ctg atc gcc gcg cgc 384

```

## RAW SEQUENCE LISTING

DATE: 06/29/2001

PATENT APPLICATION: US/09/751,299A

TIME: 14:23:08

Input Set : A:\DIVER1440-2final.txt

Output Set: N:\CRF3\06292001\I751299A.raw

```

68 Met Gly Gln Ala Ile Phe Gly Pro Asp Gly Asp Leu Ile Ala Ala Arg
69      115      120      125
71 cgc aag ctc aag cct acc cat gcg gag cgc acc gtg ttc ggc gag gga 432
72 Arg Lys Leu Lys Pro Thr His Ala Glu Arg Thr Val Phe Gly Glu Gly
73      130      135      140
75 gac ggc agc cat ctc gcg gtg cac gat acc gcc atc ggg cgc ctc ggc 480
76 Asp Gly Ser His Leu Ala Val His Asp Thr Ala Ile Gly Arg Leu Gly
77 145      150      155      160
79 gcg ctc tgt tgc tgg gag cac atc cag cca ttg tcg aaa tac gcc atg 528
80 Ala Leu Cys Cys Trp Glu His Ile Gln Pro Leu Ser Lys Tyr Ala Met
81      165      170      175
83 tac gcc gcc gac gaa cag gtc cac gtc gcg tcg tgg ccg agc ttc agc 576
84 Tyr Ala Ala Asp Glu Gln Val His Val Ala Ser Trp Pro Ser Phe Ser
85      180      185      190
87 ctc tat cgc ggc atg gcc tat gcg ctc gga ccg gag gtc aat acc gcc 624
88 Leu Tyr Arg Gly Met Ala Tyr Ala Leu Gly Pro Glu Val Asn Thr Ala
89      195      200      205
91 gca agc cag atc tac gcg gtc gag ggc ggc tgc tac gtg ctg gcg tcg 672
92 Ala Ser Gln Ile Tyr Ala Val Glu Gly Gly Cys Tyr Val Leu Ala Ser
93      210      215      220
95 tgc gcg acc gtt tcg ccg gag atg atc aag gta ttg gtg gat acg ccc 720
96 Cys Ala Thr Val Ser Pro Glu Met Ile Lys Val Leu Val Asp Thr Pro
97 225      230      235      240
99 gac aag gag atg ttc ctc aag gcc ggc ggc ggt ttt gcc atg att ttc 768
100 Asp Lys Glu Met Phe Leu Lys Ala Gly Gly Gly Phe Ala Met Ile Phe
101      245      250      255
103 ggg ccc gac ggc cgc gcc ctg gcc gag ccg ctc ccg gag acc gaa gag 816
104 Gly Pro Asp Gly Arg Ala Leu Ala Glu Pro Leu Pro Glu Thr Glu Glu
105      260      265      270
107 gga ctg ctg gtc gcc gat atc gac ctc ggc atg atc gcg ttg gcc aag 864
108 Gly Leu Leu Val Ala Asp Ile Asp Leu Gly Met Ile Ala Leu Ala Lys
109      275      280      285
111 gcg gcg gcc gat ccg gcg ggc cac tat tca cgg ccc gac gta acg cgg 912
112 Ala Ala Ala Asp Pro Ala Gly His Tyr Ser Arg Pro Asp Val Thr Arg
113      290      295      300
115 ctg ctg ctg gat cga cgt ccg gcc caa cgc gtc gtc acg ctt gat gcc 960
116 Leu Leu Leu Asp Arg Arg Pro Ala Gln Arg Val Val Thr Leu Asp Ala
117 305      310      315      320
119 gca ttc gaa ccg caa aac gag gac aag ggc gac gcg ccc gcg ctg cgc 1008
120 Ala Phe Glu Pro Gln Asn Glu Asp Lys Gly Asp Ala Pro Ala Leu Arg
121      325      330      335
123 gtg gtg gcg gaa agc gcc gcc gcc gcg cag tag 1041
124 Val Val Ala Glu Ser Ala Ala Ala Ala Gln
125      340      345
128 <210> SEQ ID NO: 2
129 <211> LENGTH: 346
130 <212> TYPE: PRT
131 <213> ORGANISM: Unknown Organism

```

W--&gt; 132 &lt;220&gt; FEATURE:

## RAW SEQUENCE LISTING

DATE: 06/29/2001

PATENT APPLICATION: US/09/751,299A

TIME: 14:23:08

Input Set : A:\DIVER1440-2final.txt

Output Set: N:\CRF3\06292001\I751299A.raw

132 &lt;223&gt; OTHER INFORMATION: Description of Unknown Organism: Obtained from an

135 &lt;400&gt; SEQUENCE: 2

136 Met Ser Glu Pro Met Thr Lys Tyr Arg Gly Ala Ala Val Gln Ala Ala

137 1 5 10 15

138 Pro Val Phe Leu Asp Leu Asp Arg Thr Val Glu Lys Ala Ile Gly Leu

139 20 25 30

140 Ile Glu Gln Ala Ala Lys Gln Asp Val Arg Leu Ile Ala Phe Pro Glu

141 35 40 45

142 Thr Trp Ile Pro Gly Tyr Pro Phe Trp Ile Trp Leu Gly Ala Pro Ala

143 50 55 60

144 Trp Gly Met Arg Phe Val Gln Arg Tyr Phe Glu Asn Ser Leu Val Arg

145 65 70 75 80

146 Gly Ser Lys Gln Trp Gln Ala Leu Ala Asp Ala Ala Arg Arg His Gly

147 85 90 95

148 Met His Val Val Ala Gly Tyr Ser Glu Arg Ala Gly Gly Ser Leu Tyr

149 100 105 110

150 Met Gly Gln Ala Ile Phe Gly Pro Asp Gly Asp Leu Ile Ala Ala Arg

151 115 120 125

152 Arg Lys Leu Lys Pro Thr His Ala Glu Arg Thr Val Phe Gly Glu Gly

153 130 135 140

154 Asp Gly Ser His Leu Ala Val His Asp Thr Ala Ile Gly Arg Leu Gly

155 145 150 155 160

156 Ala Leu Cys Cys Trp Glu His Ile Gln Pro Leu Ser Lys Tyr Ala Met

157 165 170 175

158 Tyr Ala Ala Asp Glu Gln Val His Val Ala Ser Trp Pro Ser Phe Ser

159 180 185 190

160 Leu Tyr Arg Gly Met Ala Tyr Ala Leu Gly Pro Glu Val Asn Thr Ala

161 195 200 205

162 Ala Ser Gln Ile Tyr Ala Val Glu Gly Gly Cys Tyr Val Leu Ala Ser

163 210 215 220

164 Cys Ala Thr Val Ser Pro Glu Met Ile Lys Val Leu Val Asp Thr Pro

165 225 230 235 240

166 Asp Lys Glu Met Phe Leu Lys Ala Gly Gly Gly Phe Ala Met Ile Phe

167 245 250 255

168 Gly Pro Asp Gly Arg Ala Leu Ala Glu Pro Leu Pro Glu Thr Glu Glu

169 260 265 270

170 Gly Leu Leu Val Ala Asp Ile Asp Leu Gly Met Ile Ala Leu Ala Lys

171 275 280 285

172 Ala Ala Ala Asp Pro Ala Gly His Tyr Ser Arg Pro Asp Val Thr Arg

173 290 295 300

174 Leu Leu Leu Asp Arg Arg Pro Ala Gln Arg Val Val Thr Leu Asp Ala

175 305 310 315 320

176 Ala Phe Glu Pro Gln Asn Glu Asp Lys Gly Asp Ala Pro Ala Leu Arg

177 325 330 335

178 Val Val Ala Glu Ser Ala Ala Ala Ala Gln

179 340 345

183 &lt;210&gt; SEQ ID NO: 3

184 &lt;211&gt; LENGTH: 1014

185 &lt;212&gt; TYPE: DNA

## RAW SEQUENCE LISTING

DATE: 06/29/2001

PATENT APPLICATION: US/09/751,299A

TIME: 14:23:08

Input Set : A:\DIVER1440-2final.txt

Output Set: N:\CRF3\06292001\I751299A.raw

```

186 <213> ORGANISM: Unknown Organism
188 <220> FEATURE:
189 <223> OTHER INFORMATION: Description of Unknown Organism: Obtained from an
190     environmental sample
192 <220> FEATURE:
193 <221> NAME/KEY: CDS
194 <222> LOCATION: (1)..(1014)
196 <400> SEQUENCE: 3
197 atg aaa gaa gct atc aag gtc gcc tgc gtg caa gcc gcc ccg atc tac      48
198 Met Lys Glu Ala Ile Lys Val Ala Cys Val Gln Ala Ala Pro Ile Tyr
199   1           5           10           15
201 atg gat ttg gag gcg acg gtg gac aaa acc att gag ttg atg gaa gaa      96
202 Met Asp Leu Glu Ala Thr Val Asp Lys Thr Ile Glu Leu Met Glu Glu
203           20           25           30
205 gca gca cgt aat aat gct cgt ctg atc gcc ttt ccg gaa act tgg att      144
206 Ala Ala Arg Asn Asn Ala Arg Leu Ile Ala Phe Pro Glu Thr Trp Ile
207           35           40           45
209 cca ggc tac cca tgg ttt ctt tgg ctt gac tca cca gca tgg gca atg      192
210 Pro Gly Tyr Pro Trp Phe Leu Trp Leu Asp Ser Pro Ala Trp Ala Met
211           50           55           60
213 caa ttt gta cgc caa tac cat gag aac tca ttg gag ttg gat ggc cct      240
214 Gln Phe Val Arg Gln Tyr His Glu Asn Ser Leu Glu Leu Asp Gly Pro
215   65           70           75           80
217 caa gct aag cgc att tca gat gca gcc aag cgg ttg gga atc atg gtc      288
218 Gln Ala Lys Arg Ile Ser Asp Ala Ala Lys Arg Leu Gly Ile Met Val
219           85           90           95
221 acc ctg ggg atg agt gaa cgg gtc ggt ggc acc ctt tac atc agt cag      336
222 Thr Leu Gly Met Ser Glu Arg Val Gly Gly Thr Leu Tyr Ile Ser Gln
223           100          105          110
225 tgg ttc ata ggc gat aat ggt gac acc att ggg gcc cgg cga aag ttg      384
226 Trp Phe Ile Gly Asp Asn Gly Asp Thr Ile Gly Ala Arg Arg Lys Leu
227           115          120          125
229 aaa cct act ttt gtt gaa cgt act ttg ttc ggc gaa ggg gat ggt tca      432
230 Lys Pro Thr Phe Val Glu Arg Thr Leu Phe Gly Glu Gly Asp Gly Ser
231           130          135          140
233 tcg cta gcg gtt ttc gag acg tct gtt gga agg ctg ggt ggc tta tgc      480
234 Ser Leu Ala Val Phe Glu Thr Ser Val Gly Arg Leu Gly Gly Leu Cys
235  145          150          155          160
237 tgt tgg gag cac ctt caa ccg cta aca aaa tac gct ttg tat gca caa      528
238 Cys Trp Glu His Leu Gln Pro Leu Thr Lys Tyr Ala Leu Tyr Ala Gln
239           165          170          175
241 aat gaa gag att cat tgt gcg gct tgg ccg agc ttt agc ctt tat cct      576
242 Asn Glu Glu Ile His Cys Ala Ala Trp Pro Ser Phe Ser Leu Tyr Pro
243           180          185          190
245 aat gcg gcg aaa gcc ctg ggg cct gat gtc aat gta gcg gcc tct cga      624
246 Asn Ala Ala Lys Ala Leu Gly Pro Asp Val Asn Val Ala Ala Ser Arg
247           195          200          205
249 atc tat gcc gtt gaa ggg caa tgc ttc gta cta gcg tcg tgt gcg ctc      672
250 Ile Tyr Ala Val Glu Gly Gln Cys Phe Val Leu Ala Ser Cys Ala Leu

```

## RAW SEQUENCE LISTING

DATE: 06/29/2001

PATENT APPLICATION: US/09/751,299A

TIME: 14:23:08

Input Set : A:\DIVER1440-2final.txt

Output Set: N:\CRF3\06292001\I751299A.raw

```

251      210      215      220
253 gtt tca caa tcc atg atc gat atg ctt tgt aca gat gac gaa aag cat 720
254 Val Ser Gln Ser Met Ile Asp Met Leu Cys Thr Asp Asp Glu Lys His
255 225      230      235      240
257 gcg ttg ctt ctg gct ggt gga cac tca cgt atc ata ggg cct gat 768
258 Ala Leu Leu Leu Ala Gly Gly Gly His Ser Arg Ile Ile Gly Pro Asp
259      245      250      255
261 ggt ggt gac ttg gtc gcg cct ctt gcc gaa aat gaa gag ggt att ctc 816
262 Gly Gly Asp Leu Val Ala Pro Leu Ala Glu Asn Glu Glu Gly Ile Leu
263      260      265      270
265 tac gca aac ctt gat cct gga gta cgc atc ctt gct aaa atg gcg gca 864
266 Tyr Ala Asn Leu Asp Pro Gly Val Arg Ile Leu Ala Lys Met Ala Ala
267      275      280      285
269 gac cct gct ggt cat tat tcc cgt ccc gac att act cgc ttg cta ata 912
270 Asp Pro Ala Gly His Tyr Ser Arg Pro Asp Ile Thr Arg Leu Leu Ile
271      290      295      300
273 gat cgc agc cct aaa tta ccg gta gtt gaa att gaa ggt gat ctt cgt 960
274 Asp Arg Ser Pro Lys Leu Pro Val Val Glu Ile Glu Gly Asp Leu Arg
275 305      310      315      320
277 cct tac gct ttg ggt aaa gcg tct gag acg ggt gcg caa ctc gaa gaa 1008
278 Pro Tyr Ala Leu Gly Lys Ala Ser Glu Thr Gly Ala Gln Leu Glu Glu
279      325      330      335
281 att tga 1014
282 Ile
285 <210> SEQ ID NO: 4
286 <211> LENGTH: 337
287 <212> TYPE: PRT
288 <213> ORGANISM: Unknown Organism
W--> 289 <220> FEATURE:
289 <223> OTHER INFORMATION: Description of Unknown Organism: Obtained from an
292 <400> SEQUENCE: 4
293 Met Lys Glu Ala Ile Lys Val Ala Cys Val Gln Ala Ala Pro Ile Tyr
294 1 5 10 15
295 Met Asp Leu Glu Ala Thr Val Asp Lys Thr Ile Glu Leu Met Glu Glu
296 20 25 30
297 Ala Ala Arg Asn Asn Ala Arg Leu Ile Ala Phe Pro Glu Thr Trp Ile
298 35 40 45
299 Pro Gly Tyr Pro Trp Phe Leu Trp Leu Asp Ser Pro Ala Trp Ala Met
300 50 55 60
301 Gln Phe Val Arg Gln Tyr His Glu Asn Ser Leu Glu Leu Asp Gly Pro
302 65 70 75 80
303 Gln Ala Lys Arg Ile Ser Asp Ala Ala Lys Arg Leu Gly Ile Met Val
304 85 90 95
305 Thr Leu Gly Met Ser Glu Arg Val Gly Gly Thr Leu Tyr Ile Ser Gln
306 100 105 110
307 Trp Phe Ile Gly Asp Asn Gly Asp Thr Ile Gly Ala Arg Arg Lys Leu
308 115 120 125
309 Lys Pro Thr Phe Val Glu Arg Thr Leu Phe Gly Glu Gly Asp Gly Ser
310 130 135 140

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/751,299A

DATE: 06/29/2001

TIME: 14:23:09

Input Set : A:\DIVER1440-2final.txt

Output Set: N:\CRF3\06292001\I751299A.raw

L:132 M:258 W: Mandatory Feature missing, <220> FEATURE:

L:289 M:258 W: Mandatory Feature missing, <220> FEATURE:

**STATISTICS SUMMARY**

PATENT APPLICATION: US/09/751,299A

DATE: 06/29/2001

TIME: 14:23:09

Input Set : A:\DIVER1440-2final.txt

Output Set: N:\CRF3\06292001\I751299A.raw

Application Serial Number: US/09/751,299A

Alpha or Numeric: Numeric

Application Class:

Application File Date: 12-28-2000

Art Unit: OIPE

Software Application: PatentIn

Total Number of Sequences: 4

Total Nucleotides: 2055

Total Amino Acids: 683

Number of Errors: 0

Number of Warnings: 2

Number of Corrections: 0

**MESSAGE SUMMARY**

258 W: 2 (Mandatory Feature missing)



OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/751,299A

DATE: 05/31/2001

TIME: 17:55:10

Input Set : A:\DIVER1440-2final.txt

Output Set: C:\CRF3\05312001\I751299.raw

Does Not Comply  
Corrected Diskette Needed  
pp. 3,5

C--> 3 <110> APPLICANT: Madden, Mark  
 4 Weiner, David P.  
 5 Chaplin, Jennifer A.  
 7 <120> TITLE OF INVENTION: METHODS FOR PRODUCING ENANTIOMERICALLY PURE  
 8 ALPHA-SUBSTITUTED CARBOXYLIC ACIDS  
 10 <130> FILE REFERENCE: DIVER1440-2  
 12 <140> CURRENT APPLICATION NUMBER: US 09/751,299  
 13 <141> CURRENT FILING DATE: 2001-05-01  
 15 <150> PRIOR APPLICATION NUMBER: 60/254,414  
 16 <151> PRIOR FILING DATE: 2000-12-07  
 18 <150> PRIOR APPLICATION NUMBER: 60/173,609  
 19 <151> PRIOR FILING DATE: 1999-12-29  
 21 <160> NUMBER OF SEQ ID NOS: 4  
 23 <170> SOFTWARE: PatentIn Ver. 2.1  
 25 <210> SEQ ID NO: 1  
 26 <211> LENGTH: 1041  
 27 <212> TYPE: DNA  
 28 <213> ORGANISM: Unknown Organism  
 30 <220> FEATURE:  
 31 <223> OTHER INFORMATION: Description of Unknown Organism: Obtained from an  
 32 environmental sample  
 34 <220> FEATURE:  
 35 <221> NAME/KEY: CDS  
 36 <222> LOCATION: (1)..(1041)  
 38 <400> SEQUENCE: 1  
 39 atg tcg gag ccc atg acg aag tat cgc ggc gcg gcg gtg cag gcc gcg 48  
 40 Met Ser Glu Pro Met Thr Lys Tyr Arg Gly Ala Ala Val Gln Ala Ala  
 41 1 5 10 15  
 43 ccg gtg ttc ctc gat ctc gac cgc aca gtc gag aaa gcg atc ggc ctg 96  
 44 Pro Val Phe Leu Asp Leu Asp Arg Thr Val Glu Lys Ala Ile Gly Leu  
 45 20 25 30  
 47 atc gag cag gcg gcc aag cag gac gtg cgc ctg atc gca ttc cca gag 144  
 48 Ile Glu Gln Ala Ala Lys Gln Asp Val Arg Leu Ile Ala Phe Pro Glu  
 49 35 40 45  
 51 act tgg att ccc ggc tat ccc ttt tgg ata tgg ctg ggc gcg ccg gct 192  
 52 Thr Trp Ile Pro Gly Tyr Pro Phe Trp Ile Trp Leu Gly Ala Pro Ala  
 53 50 55 60  
 55 tgg ggc atg cgc ttc gtc cag cgc tat ttc gag aat tcg ctc gtg cgc 240  
 56 Trp Gly Met Arg Phe Val Gln Arg Tyr Phe Glu Asn Ser Leu Val Arg  
 57 65 70 75 80  
 59 ggc agc aag cag tgg cag gcc ctg gcg gat gcg gcc cgc cgc cac ggc 288  
 60 Gly Ser Lys Gln Trp Gln Ala Leu Ala Asp Ala Ala Arg Arg His Gly  
 61 85 90 95  
 63 atg cat gtc gtg gcc ggc tat agc gag cgc gcg ggc ggc agc ctc tat 336  
 64 Met His Val Val Ala Gly Tyr Ser Glu Arg Ala Gly Gly Ser Leu Tyr  
 65 100 105 110  
 67 atg ggc cag gcg atc ttc ggc ccc gat ggc gat ctg atc gcc gcg cgc 384

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/751,299

DATE: 05/31/2001

TIME: 17:55:10

Input Set : A:\DIVER1440-2final.txt

Output Set: C:\CRF3\05312001\I751299.raw

```

68 Met Gly Gln Ala Ile Phe Gly Pro Asp Gly Asp Leu Ile Ala Ala Arg
69      115      120      125
71 cgc aag ctc aag cct acc cat gcg gag cgc acc gtg ttc ggc gag gga 432
72 Arg Lys Leu Lys Pro Thr His Ala Glu Arg Thr Val Phe Gly Glu Gly
73      130      135      140
75 gac ggc agc cat ctc gcg gtg cac gat acc gcc atc ggg cgc ctc ggc 480
76 Asp Gly Ser His Leu Ala Val His Asp Thr Ala Ile Gly Arg Leu Gly
77 145      150      155      160
79 gcg ctc tgt tgc tgg gag cac atc cag cca ttg tcg aaa tac gcc atg 528
80 Ala Leu Cys Cys Trp Glu His Ile Gln Pro Leu Ser Lys Tyr Ala Met
81      165      170      175
83 tac gcc gcc gac gaa cag gtc cac gtc gcg tcg tgg ccg agc ttc agc 576
84 Tyr Ala Ala Asp Glu Gln Val His Val Ala Ser Trp Pro Ser Phe Ser
85      180      185      190
87 ctc tat cgc ggc atg gcc tat gcg ctc gga ccg gag gtc aat acc gcc 624
88 Leu Tyr Arg Gly Met Ala Tyr Ala Leu Gly Pro Glu Val Asn Thr Ala
89      195      200      205
91 gca agc cag atc tac gcg gtc gag ggc ggc tgc tac gtg ctg gcg tcg 672
92 Ala Ser Gln Ile Tyr Ala Val Glu Gly Gly Cys Tyr Val Leu Ala Ser
93      210      215      220
95 tgc gcg acc gtt tcg ccg gag atg atc aag gta ttg gtg gat acg ccc 720
96 Cys Ala Thr Val Ser Pro Glu Met Ile Lys Val Leu Val Asp Thr Pro
97 225      230      235      240
99 gac aag gag atg ttc ctc aag gcc ggc ggc ggt ttt gcc atg att ttc 768
100 Asp Lys Glu Met Phe Leu Lys Ala Gly Gly Gly Phe Ala Met Ile Phe
101      245      250      255
103 ggg ccc gac ggc cgc gcc ctg gcc gag ccg ctc ccg gag acc gaa gag 816
104 Gly Pro Asp Gly Arg Ala Leu Ala Glu Pro Leu Pro Glu Thr Glu Glu
105      260      265      270
107 gga ctg ctg gtc gcc gat atc gac ctc ggc atg atc gcg ttg gcc aag 864
108 Gly Leu Leu Val Ala Asp Ile Asp Leu Gly Met Ile Ala Leu Ala Lys
109      275      280      285
111 gcg gcg gcc gat ccg gcg ggc cac tat tca cgg ccc gac gta acg cgg 912
112 Ala Ala Ala Asp Pro Ala Gly His Tyr Ser Arg Pro Asp Val Thr Arg
113      290      295      300
115 ctg ctg ctg gat cga cgt ccg gcc caa cgc gtc gtc acg ctt gat gcc 960
116 Leu Leu Leu Asp Arg Arg Pro Ala Gln Arg Val Val Thr Leu Asp Ala
117 305      310      315      320
119 gca ttc gaa ccg caa aac gag gac aag ggc gac gcg ccc gcg ctg cgc 1008
120 Ala Phe Glu Pro Gln Asn Glu Asp Lys Gly Asp Ala Pro Ala Leu Arg
121      325      330      335
123 gtg gtg gcg gaa agc gcc gcc gcc gcg cag tag' 1041
124 Val Val Ala Glu Ser Ala Ala Ala Ala Gln
125      340      345

```

128 &lt;210&gt; SEQ ID NO: 2

129 &lt;211&gt; LENGTH: 346

130 &lt;212&gt; TYPE: PRT

131 &lt;213&gt; ORGANISM: Unknown Organism

W--&gt; 132 &lt;220&gt; FEATURE:

→ See next page.

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/751,299

DATE: 05/31/2001

TIME: 17:55:10

Input Set : A:\DIVER1440-2final.txt

Output Set: C:\CRF3\05312001\I751299.raw

132 &lt;223&gt; OTHER INFORMATION: Description of Unknown Organism: Obtained from an

135 &lt;400&gt; SEQUENCE: 2

```

136 Met Ser Glu Pro Met Thr Lys Tyr Arg Gly Ala Ala Val Gln Ala Ala
137   1      5      10      15
138 Pro Val Phe Leu Asp Leu Asp Arg Thr Val Glu Lys Ala Ile Gly Leu
139           20      25      30
140 Ile Glu Gln Ala Ala Lys Gln Asp Val Arg Leu Ile Ala Phe Pro Glu
141           35      40      45
142 Thr Trp Ile Pro Gly Tyr Pro Phe Trp Ile Trp Leu Gly Ala Pro Ala
143           50      55      60
144 Trp Gly Met Arg Phe Val Gln Arg Tyr Phe Glu Asn Ser Leu Val Arg
145           65      70      75      80
146 Gly Ser Lys Gln Trp Gln Ala Leu Ala Asp Ala Ala Arg Arg His Gly
147           85      90      95
148 Met His Val Val Ala Gly Tyr Ser Glu Arg Ala Gly Gly Ser Leu Tyr
149           100     105     110
150 Met Gly Gln Ala Ile Phe Gly Pro Asp Gly Asp Leu Ile Ala Ala Arg
151           115     120     125
152 Arg Lys Leu Lys Pro Thr His Ala Glu Arg Thr Val Phe Gly Glu Gly
153           130     135     140
154 Asp Gly Ser His Leu Ala Val His Asp Thr Ala Ile Gly Arg Leu Gly
155 145           150     155     160
156 Ala Leu Cys Cys Trp Glu His Ile Gln Pro Leu Ser Lys Tyr Ala Met
157           165     170     175
158 Tyr Ala Ala Asp Glu Gln Val His Val Ala Ser Trp Pro Ser Phe Ser
159           180     185     190
160 Leu Tyr Arg Gly Met Ala Tyr Ala Leu Gly Pro Glu Val Asn Thr Ala
161           195     200     205
162 Ala Ser Gln Ile Tyr Ala Val Glu Gly Gly Cys Tyr Val Leu Ala Ser
163           210     215     220
164 Cys Ala Thr Val Ser Pro Glu Met Ile Lys Val Leu Val Asp Thr Pro
165 225           230     235     240
166 Asp Lys Glu Met Phe Leu Lys Ala Gly Gly Gly Phe Ala Met Ile Phe
167           245     250     255
168 Gly Pro Asp Gly Arg Ala Leu Ala Glu Pro Leu Pro Glu Thr Glu Glu
169           260     265     270
170 Gly Leu Leu Val Ala Asp Ile Asp Leu Gly Met Ile Ala Leu Ala Lys
171           275     280     285
172 Ala Ala Ala Asp Pro Ala Gly His Tyr Ser Arg Pro Asp Val Thr Arg
173           290     295     300
174 Leu Leu Leu Asp Arg Arg Pro Ala Gln Arg Val Val Thr Leu Asp Ala
175 305           310     315     320
176 Ala Phe Glu Pro Gln Asn Glu Asp Lys Gly Asp Ala Pro Ala Leu Arg
177           325     330     335
178 Val Val Ala Glu Ser Ala Ala Ala Ala Gln
179           340     345

```

183 &lt;210&gt; SEQ ID NO: 3

184 &lt;211&gt; LENGTH: 1014

185 &lt;212&gt; TYPE: DNA

*Description  
cut off due  
to missing  
<220> feature.*

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/751,299

DATE: 05/31/2001

TIME: 17:55:10

Input Set : A:\DIVER1440-2final.txt

Output Set: C:\CRF3\05312001\I751299.raw

```

186 <213> ORGANISM: Unknown Organism
188 <220> FEATURE:
189 <223> OTHER INFORMATION: Description of Unknown Organism: Obtained from an
190     environmental sample
192 <220> FEATURE:
193 <221> NAME/KEY: CDS
194 <222> LOCATION: (1)..(1014)
196 <400> SEQUENCE: 3
197 atg aaa gaa gct atc aag gtc gcc tgc gtg caa gcc gcc ccg atc tac 48
198 Met Lys Glu Ala Ile Lys Val Ala Cys Val Gln Ala Ala Pro Ile Tyr
199 1 5 10 15
201 atg gat ttg gag gcg acg gtg gac aaa acc att gag ttg atg gaa gaa 96
202 Met Asp Leu Glu Ala Thr Val Asp Lys Thr Ile Glu Leu Met Glu Glu
203 20 25 30
205 gca gca cgt aat aat gct cgt ctg atc gcc ttt ccg gaa act tgg att 144
206 Ala Ala Arg Asn Asn Ala Arg Leu Ile Ala Phe Pro Glu Thr Trp Ile
207 35 40 45
209 cca ggc tac cca tgg ttt ctt tgg ctt gac tca cca gca tgg gca atg 192
210 Pro Gly Tyr Pro Trp Phe Leu Trp Leu Asp Ser Pro Ala Trp Ala Met
211 50 55 60
213 caa ttt gta cgc caa tac cat gag aac tca ttg gag ttg gat ggc cct 240
214 Gln Phe Val Arg Gln Tyr His Glu Asn Ser Leu Glu Leu Asp Gly Pro
215 65 70 75 80
217 caa gct aag cgc att tca gat gca gcc aag cgg ttg gga atc atg gtc 288
218 Gln Ala Lys Arg Ile Ser Asp Ala Ala Lys Arg Leu Gly Ile Met Val
219 85 90 95
221 acc ctg ggg atg agt gaa cgg gtc ggt ggc acc ctt tac atc agt cag 336
222 Thr Leu Gly Met Ser Glu Arg Val Gly Gly Thr Leu Tyr Ile Ser Gln
223 100 105 110
225 tgg ttc ata ggc gat aat ggt gac acc att ggg gcc cgg cga aag ttg 384
226 Trp Phe Ile Gly Asp Asn Gly Asp Thr Ile Gly Ala Arg Arg Lys Leu
227 115 120 125
229 aaa cct act ttt gtt gaa cgt act ttg ttc ggc gaa ggg gat ggt tca 432
230 Lys Pro Thr Phe Val Glu Arg Thr Leu Phe Gly Glu Gly Asp Gly Ser
231 130 135 140
233 tcg cta gcg gtt ttc gag acg tct gtt gga agg ctg ggt ggc tta tgc 480
234 Ser Leu Ala Val Phe Glu Thr Ser Val Gly Arg Leu Gly Gly Leu Cys
235 145 150 155 160
237 tgt tgg gag cac ctt caa ccg cta aca aaa tac gct ttg tat gca caa 528
238 Cys Trp Glu His Leu Gln Pro Leu Thr Lys Tyr Ala Leu Tyr Ala Gln
239 165 170 175
241 aat gaa gag att cat tgt gcg gct tgg ccg agc ttt agc ctt tat cct 576
242 Asn Glu Glu Ile His Cys Ala Ala Trp Pro Ser Phe Ser Leu Tyr Pro
243 180 185 190
245 aat gcg gcg aaa gcc ctg ggg cct gat gtc aat gta gcg gcc tct cga 624
246 Asn Ala Ala Lys Ala Leu Gly Pro Asp Val Asn Val Ala Ala Ser Arg
247 195 200 205
249 atc tat gcc gtt gaa ggg caa tgc ttc gta cta gcg tcg tgt gcg ctc 672
250 Ile Tyr Ala Val Glu Gly Gln Cys Phe Val Leu Ala Ser Cys Ala Leu

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/751,299

DATE: 05/31/2001

TIME: 17:55:10

Input Set : A:\DIVER1440-2final.txt

Output Set: C:\CRF3\05312001\I751299.raw

```

251      210      215      220
253 gtt tca caa tcc atg atc gat atg ctt tgt aca gat gac gaa aag cat 720
254 Val Ser Gln Ser Met Ile Asp Met Leu Cys Thr Asp Asp Glu Lys His
255 225      230      235      240
257 gcg ttg ctt ctg gct ggt ggt gga cac tca cgt atc ata ggg cct gat 768
258 Ala Leu Leu Leu Ala Gly Gly Gly His Ser Arg Ile Ile Gly Pro Asp
259      245      250      255
261 ggt ggt gac ttg gtc gcg cct ctt gcc gaa aat gaa gag ggt att ctc 816
262 Gly Gly Asp Leu Val Ala Pro Leu Ala Glu Asn Glu Glu Gly Ile Leu
263      260      265      270
265 tac gca aac ctt gat cct gga gta cgc atc ctt gct aaa atg gcg gca 864
266 Tyr Ala Asn Leu Asp Pro Gly Val Arg Ile Leu Ala Lys Met Ala Ala
267      275      280      285
269 gac cct gct ggt cat tat tcc cgt ccc gac att act cgc ttg cta ata 912
270 Asp Pro Ala Gly His Tyr Ser Arg Pro Asp Ile Thr Arg Leu Leu Ile
271      290      295      300
273 gat cgc agc cct aaa tta ccg gta gtt gaa att gaa ggt gat ctt cgt 960
274 Asp Arg Ser Pro Lys Leu Pro Val Val Glu Ile Glu Gly Asp Leu Arg
275 305      310      315      320
277 cct tac gct ttg ggt aaa gcg tct gag acg ggt gcg caa ctc gaa gaa 1008
278 Pro Tyr Ala Leu Gly Lys Ala Ser Glu Thr Gly Ala Gln Leu Glu Glu
279      325      330      335
281 att tga' 1014
282 Ile
285 <210> SEQ ID NO: 4
286 <211> LENGTH: 337
287 <212> TYPE: PRT
288 <213> ORGANISM: Unknown Organism
W--> 289 <220> FEATURE:
289 <223> OTHER INFORMATION: Description of Unknown Organism: Obtained from an
292 <400> SEQUENCE: 4
293 Met Lys Glu Ala Ile Lys Val Ala Cys Val Gln Ala Ala Pro Ile Tyr
294 1 5 10 15
295 Met Asp Leu Glu Ala Thr Val Asp Lys Thr Ile Glu Leu Met Glu Glu
296 20 25 30
297 Ala Ala Arg Asn Asn Ala Arg Leu Ile Ala Phe Pro Glu Thr Trp Ile
298 35 40 45
299 Pro Gly Tyr Pro Trp Phe Leu Trp Leu Asp Ser Pro Ala Trp Ala Met
300 50 55 60
301 Gln Phe Val Arg Gln Tyr His Glu Asn Ser Leu Glu Leu Asp Gly Pro
302 65 70 75 80
303 Gln Ala Lys Arg Ile Ser Asp Ala Ala Lys Arg Leu Gly Ile Met Val
304 85 90 95
305 Thr Leu Gly Met Ser Glu Arg Val Gly Gly Thr Leu Tyr Ile Ser Gln
306 100 105 110
307 Trp Phe Ile Gly Asp Asn Gly Asp Thr Ile Gly Ala Arg Arg Lys Leu
308 115 120 125
309 Lys Pro Thr Phe Val Glu Arg Thr Leu Phe Gly Glu Gly Asp Gly Ser
310 130 135 140

```

*Description cut off due  
to missing <220> feature.*

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/09/751,299

DATE: 05/31/2001

TIME: 17:55:11

Input Set : A:\DIVER1440-2final.txt

Output Set: C:\CRF3\05312001\I751299.raw

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:132 M:258 W: Mandatory Feature missing, <220> FEATURE:

L:289 M:258 W: Mandatory Feature missing, <220> FEATURE: